

# Lifespan variation by causes of death & Gender Gap in Europe

Data & Methods Assignment

Jesús Daniel Zazueta Borboa

European Doctoral School of Demography

## 1 Introduction

The topic of this research is lifespan variation and its gender gap by cause of death in Europe for the period 1980-2015. Lifespan variation, also referred to as lifespan inequality, captures the heterogeneity at age at death, while life expectancy provides an average of the length of life. Both are useful indicators of population health and performances of the health system ([van Raalte et al., 2018](#); [Aburto et al., 2018](#)). The objective of this research is to assess the contribution of different causes of deaths to the trend of lifespan variation and to the gender gap, and add to the literature in this topic. This document describes the data for this research.

## 2 Data

To measure inequality in the length of life, it is indispensable to have life table data. But, in order to take into account the contribution of causes of death, it is necessary to combine period life tables with an external data source. In this research, period life tables were retrieved from the Human Mortality Database (HMD) for the period 1980 to 2015, while cause of death data come from the World Health Organization (WHO) Mortality Database.

There are two primary reasons to justify the use of both data sources. First, HMD, and WHO are recognized as high-quality public data sources to analyze mortality and to make cross-national comparisons. Second, previous research on this topic, often combine both data sources ([Aburto et al., 2018](#); [Seligman et al., 2016](#)) or use WHO mortality data

as a primary source ([Alvarez et al., 2020](#); [Cao et al., 2017](#); [Nau and Firebaugh, 2012](#)). In that way my results will be easily comparable with the rest of the literature.

## 2.1 Human Mortality Database

The HMD is a joint project between the University of California at Berkeley and the Max Planck Institute for Demographic Research ([HMD, 2018](#)). Most of the countries in the HMD data are European, and for some countries such as France, Sweden, and Denmark, period and cohort life tables are available going back to the XVIII century. Other non-European countries are included, such as the United States of America, Canada, Chile, Australia, New Zealand, Israel, Japan, Taiwan, Republic of Korea and Hong Kong.

As a database, HMD provides the mid-year-population, number of deaths, number of birth, deaths in the Lexis Diagram, and period and cohort life tables. The key features of HMD are the availability of the information on the methods used to construct these life tables, and a special appendix for each country, which describes the treatment and the source of the data. In this research, we use period life table by sex and single year of age (0 -110+) for the period 1980 to 2015 for 22 countries<sup>1</sup>. For those selected countries, the data is collected from the National Statistic Office, and used by the HMD team to estimate the life tables.

## 2.2 World Health Organization

The WHO Mortality Data contains the number of deaths by country, year, sex, age groups, and cause of death as far back as 1950. Data are included only for countries reporting data coded according to the International Classification of Diseases (ICD) ([WHO, 2019](#)). For this analysis, we included the same selected countries as for the HMD and for the same time framework.

One of the limitations of this database is that the number of deaths is in five-years interval age groups except for death at age 0 and 1. Another limitation is that for some periods and countries, the last age group varies between 80+ or 85+. In order to fit the data with the 1-year age groups of the HMD and increase the accuracy of the information,

---

<sup>1</sup>Austria, Belgium, Bulgaria, Czech Republic, Denmark, Estonia, Finland, France, Ireland, Italy, Latvia, Lithuania, Netherlands, Norway, Portugal, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom, Ukraine

I will ungroup and smooth the deaths by each cause of death using efficient estimation of smoothed distribution using the R packages *ungroup* (Rizzi et al., 2015, 2016).

### 2.2.1 Cause of death classification

For the period of analysis, there are three revisions of the ICD classification of causes of death, the eighth, ninth, and tenth revision. I will group the deaths in seven categories following previous research to increase accuracy and comparability (Janssen and Kunst, 2004; Aburto et al., 2018). The seven causes of death are: 1) Cancer sensitive to smoking, 2) Cancer not sensitive to smoking, 3) Cardiovascular diseases, 4) Non-infectious respiratory diseases, 5) Infectious respiratory diseases, 6) External causes and 7) Rest of causes. As an appendix of this thesis, the detailed information of ICD Classification codes will be provided.

## References

- Aburto, J. M., Wensink, M., van Raalte, A., and Lindahl-Jacobsen, R. (2018). Potential gains in life expectancy by reducing inequality of lifespans in denmark: an international comparison and cause-of-death analysis. *BMC public health*, 18(1):831.
- Alvarez, J.-A., Aburto, J. M., and Canudas-Romo, V. (2020). Latin american convergence and divergence towards the mortality profiles of developed countries. *Population Studies*, 74(1):75–92.
- Cao, B., Bray, F., Beltrán-Sánchez, H., Ginsburg, O., Soneji, S., and Soerjomataram, I. (2017). Benchmarking life expectancy and cancer mortality: global comparison with cardiovascular disease 1981-2010. *bmj*, 357:j2765.
- HMD (2018). Human mortality database. *University of California, Berkley & Max Planck Institute for Demographic Research*.
- Janssen, F. and Kunst, A. E. (2004). Icd coding changes and discontinuities in trends in cause-specific mortality in six european countries, 1950-99. *Bulletin of the World Health Organization*, 82:904–913.

- Nau, C. and Firebaugh, G. (2012). A new method for determining why length of life is more unequal in some populations than in others. *Demography*, 49(4):1207–1230.
- Rizzi, S., Gampe, J., and Eilers, P. H. (2015). Efficient estimation of smooth distributions from coarsely grouped data. *American Journal of Epidemiology*, 182(2):138–147.
- Rizzi, S., Thinggaard, M., Engholm, G., Christensen, N., Johannesen, T. B., Vaupel, J. W., and Lindahl-Jacobsen, R. (2016). Comparison of non-parametric methods for ungrouping coarsely aggregated data. *BMC medical research methodology*, 16(1):59.
- Seligman, B., Greenberg, G., and Tuljapurkar, S. (2016). Equity and length of lifespan are not the same. *Proceedings of the National Academy of Sciences*, 113(30):8420–8423.
- van Raalte, A. A., Sasson, I., and Martikainen, P. (2018). The case for monitoring life-span inequality. *Science*, 362(6418):1002–1004.
- WHO (2019). Health statistics and information systems. *World Health Organization online*. Accessed 25 March 2020.